DOCUMENT RESUME

ED 415 616 EC 306 096

AUTHOR	Liu, Kristin; Thurlow, Martha; Vieburg, Kayleen; El-Sawâf, Hamdy; Ruhland, Aaron
TITLE	Resources: Limited English Proficient Students in National and Statewide Assessments. State Assessment Series, Minnesota Report 8.
INSTITUTION	National Center on Educational Outcomes, Minneapolis, MN.; Minnesota State Dept. of Children, Families, and Learning, St. Paul.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	1996-08-00
NOTE	54p.
CONTRACT	R279A50011
AVAILABLE FROM	National Center on Educational Outcomes, University of Minnesota, 350 Elliott Hall, 75 East River Road, Minneapolis, MN 55455; telephone: 612-624-8561; fax: 612-624-0879; World Wide Web: http://www.coled. umn.edu/NCEO (document may be copied without charge; additional print copies $15).
PUB TYPE	Reference Materials - Bibliographies (131)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Academic Standards; Definitions; *Disabilities; *Educational Assessment; *Graduation Requirements; High Schools; *Limited English Speaking; *Minimum Competency Testing; Performance Tests; Psychometrics; *State Standards; Student Evaluation
IDENTIFIERS	*Minnesota; Testing Accommodations (Disabilities)

ABSTRACT

This annotated bibliography of 66 resources was developed as part of a federally funded project in Minnesota to promote the participation of students with limited English proficiency (LEP) and students with disabilities in the Minnesota Graduation Standards. Specifically, the project is encouraging participation in the Basic Standards Exams of reading, mathematics, and written composition, and in the performance-based assessments in the Profile of Learning. This bibliography lists reference materials on accommodation issues that relate to LEP students and assessment. The bibliography is based on searches of the Psychology Literature database, the ERIC database, and the World Wide Web, as well as documents from state and national agencies and libraries. The bibliography resources are listed alphabetically by author within each of four topics: (1) LEP students in assessments and educational reform; (2) general testing issues; (3) psychometric issues in assessment; and (4) miscellaneous--such as definitions of "limited English proficient" and difficulties associated with educating LEP students. Document dates range from 1978 through 1996 with the great majority published in the 1990s. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 415 616

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Resources: Limited English Proficient Students in National and Statewide Assessments

EC 306096

BEST COPY AVAILABLE

STATE ASSESSMENT SERIES
Minnesota Report 8

Resources: Limited English Proficient Students in National and Statewide Assessments

Minnesota Assessment Project

Project Staff:
Constance Anderson • Cathy Wagner • Michael Tillman
Minnesota Department of Children, Families and Learning

Project Advisors:
Karon Hergenhahn • Bounlieng Phommasouvanh • Marilyn Sorensen • Elizabeth Watkins
Minnesota Department of Children, Families and Learning

Prepared By:
Kristin Liu • Martha Thurlow • Kayleen Vieburg • Hamdy El-Sawaf • Aaron Ruhland
University of Minnesota

August 1996

The Minnesota Assessment Project is a four-year, federally funded effort awarded to the Minnesota Department of Children, Families and Learning from the U.S. Department of Education, Office of Educational Research and Improvement. The project's goal is to promote and evaluate the participation of students with limited English proficiency and students with disabilities in Minnesota's Graduation Standards. Specifically, the project will examine ways in which students with limited English and students with disabilities can participate in the Basic Standards Exams of reading, mathematics and written composition and in the performance-based assessments of the high standards in the Profile of Learning.

This project is supported, in part, by a grant to the Minnesota Department of Children, Families and Learning from the U.S. Department of Education, Office of Educational Research and Improvement (Grant #R279A50011). Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

Permission is granted to copy this document without charge. Additional print copies may be purchased for $15.00 from:

NATIONAL CENTER ON EDUCATIONAL OUTCOMES
University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612.624.8561 • Fax 612.624.0879
http://www.coled.umn.edu/NCEO

This document is available in alternative formats upon request.

Overview

In 1995 the Department of Children, Families and Learning (CFL) and the National Center on Educational Outcomes (NCEO) received a grant from the U.S. Department of Educational Research and Improvement (OERI) to evaluate the development and implementation of the Minnesota Basic Standards Tests. The Basic Standards Tests are high stakes tests that students take to receive a high school diploma. Traditionally, many limited-English-proficient (LEP) students across the country have been exempted from achievement testing because educators thought that the students were not proficient enough in English to take the test. Lacelle-Peterson and Rivera (1994, p. 70) say that while this consideration for the students is admirable, exempting LEP students creates a kind of "systemic ignorance" about their progress. Exempting the students may also create an inaccurate picture of the overall student achievement in a district and will make it difficult for educational reform to address the needs of all students. Therefore, one of the goals of CFL and NCEO is to encourage districts in Minnesota to include as many of their LEP students in the assessment as possible.

The purpose of this annotated bibliography is to provide an updated list of reference materials on accommodation issues that relate to LEP students and assessment. To find material for this bibliography the Psychology Literature database, the ERIC database, and the World Wide Web were searched. In addition to these materials, other items were obtained from state and national agencies (e.g., Center for Applied Linguistics, North Central Regional Educational Laboratory [NCREL]) as well as from area libraries and university bookstores with educational textbooks. The current search focused on documents published within the past ten years, but did include earlier resources if it was thought that they were still relevant.

When searching the Psychology Literature database and ERIC, key words such as the following were used in various combinations: assessment, ESL, bilingual, education, assessment, testing, graduation, high school, language, proficiency, racial differences, ethnic differences, minimum competency tests, and academic achievement. The search turned up few documents on accommodation issues. Therefore, this bibliography emphasizes the following four areas:

- LEP Students in Assessments and Educational Reform.
- General Testing Issues.
- Psychometric Issues in Assessment.
- Miscellaneous. An additional miscellaneous section contains resources that give definitions of "limited English proficient," discusses difficulties associated with

educating LEP students, and examines relevant problems related to raising the standards for graduation.

A future bibliography is planned that will examine critical issues of second language acquisition in the three content areas assessed on the Basic Standards Tests: reading, math, and writing. Information on second language acquisition could be useful when developing grading standards for a high school graduation test (i.e., Should LEP students be scored using a different rubric on a writing test because that skill is slower to develop in a second language than mathematical skills?).

The bibliography resources listed in this document are divided into categories by topic. This was done for several reasons: First, the issue of assessment for LEP students includes more than just state and national assessments; it also includes assessment for language proficiency, placement in a language learning program, exit from a language learning program, and assessment for special education services. Thus, an article with useful information on assessment practices may not be directly related to national and statewide assessments. Second, when designing and implementing assessments with LEP students, second language acquisition issues must be taken into consideration. Third, issues relating to government policy and national mandates for testing are useful background information even if these documents do not specifically address the topic of LEP students.

It is important to note that this bibliography is not exhaustive due to the large number of potential resources available. A second annotated bibliography is planned to cover resources that are specific to certain subtopics, such as second language acquisition and competency development in specific content areas.

LEP Students in Assessments and Educational Reform

Anstrom, K. (1996, Summer). Defining the limited-English proficient student population. *Directions in Language and Education* [On-line], *1(9)*. Available National Clearinghouse of Bilingual Education: http://www.ncbe.gwu.edu/ncbepubs/directions

This document explains the different ways in which the LEP population in the U.S. has been estimated and the many definitions of LEP that account for the different estimations. The author compares the federal definition of LEP and state definitions for California, New York, and Texas. Across these three states, little to no agreement occurred in the way LEP student status was defined. Cheung's (1994) definition of LEP from "The Feasibility of Collecting Comparable National Statistics about Students with Limited English Proficiency: A Final Report of the LEP Student Counts Study" is listed as a potential starting point for creating a standardized definition. Next, the advantages and disadvantages of having a standardized definition are discussed. Also included is an appendix with an annotated bibliography of recent LEP population estimates.

Anstrom, K., & Kindler, A. (1996). Federal policy, legislation, and education reform: The promise and the challenge for language minority students [On-line]. Available National Clearinghouse of Bilingual Education Resource Collection Series No. 5.: http://www.ncbe.gwu.edu/ncbepubs/resources/fedpol.html

This document provides "an overview of the issues and legislation pertinent to the attainment of educational equity and excellence for language minority students." It discusses the current language minority and LEP student populations and the difficulties associated with educating these populations. It describes effective bilingual programs and the teaching practices that are a part of such programs. It also lists the necessary factors for successful implementation of bilingual programs within the context of a whole-school approach to reform. Finally, it lists and describes the three pieces of legislation that are involved in the education reform movement:

- *Goals 2000: Educate America Act*, which provides funding for the reformation of the national education system.
- *Improving America's Schools Act (IASA) of 1994*, which re-authorizes the Elementary and Secondary Education Act for five more years, funds programs for disadvantaged and LEP students, and promotes implementation of professional development programs.

- *The Title II: Dwight D. Eisenhower Professional Development Program*, which provides funds for professional development. These funds can be used to improve teachers' ability to work with LEP students.

August, D., Hakuta, K., Olguin, F., & Pompa, D. (1995). LEP students and Title 1: A guidebook for educators [On-line]. Available National Clearinghouse of Bilingual Education Resource Collection Series: http://www.ncbe.gwu.edu/ncbepubs/resource/lepguide/

This document explains what the IASA of 1994 requires for LEP students and what implications the law has for them. The document was written for Title I Coordinators, administrators of programs that serve LEP students, policy makers, and parents of LEP students. Included in the discussion of implications are recommendations for ways that states and schools can facilitate compliance with the law while meeting the needs of LEP students. Areas covered in the document are:

- State plans relating to development, content standards, assessment, and support for teaching and learning.
- Local educational agency (LEA) plans.
- Schoolwide programs.
- Targeted assistance schools.
- Assessment and LEA and school improvement.
- State assistance for school support and improvement.
- Parent involvement.
- Professional development.

The authors believe that LEP students should be eligible for services on the same basis as other students served by Title I and should be held to the same high content and performance standards required of all Title I students.

August, D., & Hakuta, K. (1994). Evaluating the inclusion of L.E.P. students in systemic reform. In *Issues and strategies in evaluating systemic reform*. Papers prepared for the U.S. Department of Education, Office of the Under Secretary, Planning and Evaluation Service.

This paper discusses many important issues for ensuring that LEP students are included in the design and implementation of the local, state, and national education reforms. For each level of reform, it describes "the various arenas, benchmarks, and methodologies that comprise the evaluation." It emphasizes the importance of guaranteeing appropriate participation of LEP students in instruction and assessment, the need for continuing improvements in assessment, and the need to consider LEP students as part of a broader analysis of systemic reform instead of singling them out as a "special" group.

August, D., Hakuta, K., & Pompa, D. (1994). *For all students: Limited-English-proficient students and Goals 2000. A discussion paper.* Paper presented at the National Academy of Education Panel Meeting No. 15, Washington, DC.

This paper compiles information and documents that resulted from the following meetings:

- Stanford Working Group on Federal Education Programs for LEP Students.
- Two Washington, D.C. meetings on standards and assessment and LEP students.
- Regional meetings to discuss the implications, at both the local and state levels, of educational reform on LEP students.

The paper highlights the major issues that arise for LEP students as a result of the national education reform and makes recommendations for ways to address these issues. Recommendations fall into the following categories: standards, assessment, accountability, research and development, Native American issues, National Skill Standard Board, and language and culture.

In the appendices there are specific recommendations relating to grants for individuals who develop "voluntary national opportunity-to-learn standards."

Bilingual Education Commission. (1994). *Striving for success: The education of bilingual pupils* (The Commonwealth of Massachusetts). Boston: Bilingual Education Commission.

This report focuses on presenting a "Coherent understanding of the complex process of educating a language minority pupil" (p. 7). The report is the result of a one year study that included "Interviews with experts in the field, a review of a survey of bilingual education program directors, a review of Massachusetts Department of Education data, documents and annual reports, and a review of current literature on bilingual education theories and practices." The Bilingual Education Commission studied each component that influenced the education of a bilingual student:

- Principals, teachers, and other school personnel.
- Parents.
- Programs.
- Schools.
- State government.

After giving a brief overview of the history of bilingual education and the continued need for it, each of the five components above is discussed in detail. Finally, there is a section of recommendations for achieving high quality bilingual programs that meet the requirements of the Education Reform Act of 1993. (Note: The section on Programs includes a discussion of program assessments and high standards for LEP students.)

Bilingual Education Expert Testifies Against National Achievement Test. (1991, September 15). *The National Association for Bilingual Education News, 15(1),* 1-2.

This article briefly discusses Dr. Edward De Avila's address to the House Subcommittee on Select Education at the time when the Goals 2000 initiative was introduced. Dr. De Avila argued against a national achievement test for several reasons:

- It is unfair to groups of children who have not received the same instruction as other groups.
- A national test will strongly influence curriculum.

- An additional test may cause low-achieving students to stay away or drop out of school.

- The test is unnecessary because the information it would provide is already available.

Dr. De Avila also states that it is difficult to determine which LEP students are capable of taking tests because there is not a consistent definition of LEP, there is a large difference in the language skills that individual LEP students possess, there is a shortage of appropriate tests that are "psychometrically and linguistically sound," and there is a lack of decision-making models for the testing process.

Bond, L.A., Braskamp, D., & Roeber, E. (1996). Part I: Assessment of students with disabilities and limited-English-proficient students. *The status report of the assessment programs in the United States: State student assessment programs database school year 1994–1995*. Oak Brook, IL: North Central Regional Education Laboratories and Council of Chief State School Officers.

This chapter discusses the current status of IASA legislation to hold all students to the same high standards. It addresses two questions:

- How many students with disabilities and LEP students participate in statewide testing programs?

- What kinds of accommodations are allowed for these types of students who do participate in assessments?

The researchers found that while the majority of states have written guidelines for the participation of students with disabilities and know what the participation rates are for these students, few of them have similar information for LEP students. Few states collect specific data on the numbers of students in either category who are exempted from testing. Most states follow the recommendations in an IEP plan to determine whether a student with disabilities should participate in testing. For LEP students, the level of English proficiency and the number of years in ESL classes are the two major factors considered. Many states allow accommodations for students with disabilities who participate in testing, but only a small number of states allow accommodations for LEP students. Finally, the findings from a 1995 survey that asked assessment directors to describe Title I assessment and evaluation plans are provided.

Chamberlain, P. & Medinos-Landurand, P. (1991). Practical considerations for the assessment of LEP students with special needs. In Hamayan, E.V. & Damico, J.S. (Eds.). *Limiting bias in the assessment of bilingual students* (pp. 122-156). Austin, TX: Pro-Ed.

This lengthy chapter in a book on assessing bilingual students deals with many issues that are applicable to testing situations other than testing for special education placement. The chapter addresses cultural and linguistic factors in a student's background that may influence assessment and gives strategies for dealing with these factors. The chapter begins with a definition of culture and then looks at types of cultural insensitivity. Bias in assessment is a type of cultural insensitivity and it may happen for reasons that are related to the testing situation itself or related to larger factors such as schooling and child-raising differences between cultures. A list of thirteen cultural variables that influence the assessment process is given:

- Cultural attitudes toward competition.
- Cultural attitudes toward time.
- Cultural preferences for dealing with many tasks at once versus preferences for dealing with one task at a time (i.e., testing requires students to do one thing at a time).
- Cultural styles of body movement (i.e., some cultures emphasize a more active style which conflicts with the passive style required in a testing situation).
- Cultural attitudes toward use of space (i.e., some cultures value closeness, while a testing situation values separation and distance; a difference in the use of space may be stressful to an examinee).
- Cultural attitudes about physical contact.
- Cultural uses of eye contact (i.e., some students may not look directly at a teacher because it is the accepted style in their culture; as a result they may miss important testing instructions that are done non-verbally).
- Cultural attitudes toward gender (i.e., female examinees may feel uncomfortable with a male test administrator; male examinees may not know how to answer certain questions relating to traditionally female topics, etc.).

- Cultural orientations regarding individuals versus families (i.e., a testing situation in which individual achievement is emphasized will not motivate a student from a culture in which individual achievement is not emphasized).

- Cultural norms about verbal and non-verbal communication (e.g., a teacher may nod at a student to say "yes" and the student may interpret the nod to mean "no").

- Cultural beliefs in fate versus individual responsibility (i.e., students from certain cultures may believe there is nothing they can do to improve their test performance because that performance is predetermined by forces beyond their control).

- Cultural trends toward a certain perceptual style (i.e., testing requires students to answer questions out of context and some students may come from a culture in which students do not learn to isolate items from their context).

- Cultural trends toward a particular cognitive style (i.e., tests value an analytical, reflective style that deals with problems step-by-step; other cultures may value a more global-intuitive style where decisions are made more spontaneously).

Following the list of cultural factors is a brief discussion of linguistic factors that may influence test performance. Included in the list are language use patterns, language loss, code switching, and dialectal variance. Any one of these linguistic factors may negatively influence the testing situation if educators are unaware of them. The authors give a list of five strategies that may help overcome test bias when assessing bilingual students. The chapter ends with a section on LEP students with disabilities and what the law requires of assessment for these students.

Chamot, A.U. (1992, August). *Changing instruction for language minority students to achieve national goals* [On-line]. Paper from the Third Plenary Session of The Third National Research Symposium on Limited English Proficient Student Issues, Washington, DC. Available: http://www.ncbe.gwu.edu/ncbepubs/symposia/chamot.html

This paper identifies "some of the major academic needs of language minority students who are learning English in secondary schools" and suggests ways to meet those needs. The paper discusses the major academic needs of LEP students: language development, instructional time, subject matter concepts, learning strategies, and self-efficacy. Following is a literature review on characteristics of effective programs for LEP students and effective instructional practices

within those programs. The author cites seven characteristics of effective school programs for language minority students:

- Strong leaders who are committed to supporting LEP students.
- High attitudes and expectations for LEP students.
- Strong academic focus.
- Curriculum that is selective and is closely related to the curriculum for native English speakers.
- Expert teachers who are trained well and who receive support from the district.
- Involved parents who feel that they are welcome in the child's school.
- Academic support in the areas of native language services, extended educational opportunities, and alternatives to standardized achievement tests.

The author discusses four categories of teaching approaches and techniques that lead to higher achievement for LEP students. Based on the research cited, the author believes that instruction for LEP students needs to change in the areas of language support, instructional time, and teaching practices. The author recommends access to bilingual staff, teaching language skills across the curriculum and teaching content in the ESL class, more time spent in learning academic skills, communicating high expectations for academic achievement, and teaching learning strategies to help students become better learners.

Cloud, N. (1991). Educational assessment. In E.V. Hamayan & J.S. Damico (Eds.). (1991). *Limiting bias in the assessment of bilingual students* (pp. 219-246). Austin, TX: Pro-Ed.

This chapter from a book on assessing bilingual students addresses assessment for special education programs. However, it begins with a discussion of the impact of language proficiency and interrupted schooling on the assessment process. When assessing a student, the author recommends documenting language proficiency in both languages, as well as taking into account the student's communicative competence and cognitive academic language proficiency (CALP). The test administrator should also take into account the student's educational history. Interrupted schooling can affect a student's language use, familiarity with the educational system, role in society, self-esteem, and motivation. Immigrants and migrant workers (people who may spend part of the year in their native country and part in another

country) frequently experience school failure because of the stress, confusion, and the loss of support that they experience in the process of moving from one place to another. Other factors such as a low socioeconomic status and the level of acculturation to the U.S. can influence educational success as well. The author provides sample questionnaires that elicit the type of information that educators should take into consideration. These questionnaires ask about the home context, the current classroom context, and the student's educational history. The chapter ends with a discussion of formal and informal tests in the student's first and second language that can be used to document the need for special education services. A list of cautions and criticisms for each type of test is provided to assist the educator in selecting an appropriate assessment.

Council of Chief State School Officers. (1992a). *Recommendations for improving the assessment and monitoring of students with limited English proficiency.* Washington, DC: Council of Chief State School Officers, Resource Center on Educational Equity.

This publication encourages improvements in the way LEP students are assessed as well as in the way the data on their educational status and progress are reported and used. It is divided into four sections:

- Issues regarding the assessment of LEP students.
- Procedures for collecting data.
- A model that illustrates the recommended assessment and service process.
- Policy recommendations at the federal and state levels. It gives recommendations that include principles and ideal practices to be used in programs for LEP students.

The report specifically provides direction for improving and making more uniform LEP screening and assessment procedures for three types of assessments:

- Classification as an LEP student.
- Placement in a language learning program.
- Reclassification – monitoring academic progress and making changes in the services received.

In addition, the report contains recommendations concerning state-level-data collection efforts focused on LEP students.

Council of Chief State School Officers. (1992b). *Summary of recommendations and policy implications for improving the assessment and monitoring of students with limited English proficiency* [On-line]. Available: http://www.ccsso.org/leppol.htm

This document summarizes a larger CCSO document titled *Recommendations for Improving the Assessment and Monitoring of Students with Limited English Proficiency.* This condensed version gives a list of recommendations for improving test screening and assessment procedures for LEP students, and improving the data-collection methods used by states. First, definitions of LEP and FEP (fully English proficient) are given because data from different sites can only be compared if students are classified in the same way. Then the recommendations for improving testing are given. Recommendations are broken down into two main categories:

- Identification and placement of students.
- State and federal policy implications.

Council of Chief State School Officers. (1990). *School success for limited English proficient students: The challenge and state response.* Washington, DC: Council of Chief State School Officers, Resource Center on Educational Equity.

This report discusses a study done by CCSSO that examined how offices within a state education agency (SEA) address the needs of LEP students. The purpose of the report is to share information on what states are doing to promote efforts to develop more effective programs for LEP students. The chapters are organized into the following topics:

- Demographics – Who are LEP students?
- A summary of goals, methodology, and results of two surveys sent to SEA personnel. The surveys found that there is a lack of adequate data on access issues (i.e., enrollment) and on "substantive educational issues" (i.e., academic status of LEP students) that prevent SEAs from developing effective programs to serve their LEP students.

- SEA initiatives to serve LEP children in California, Illinois, New Jersey, and Texas. This section includes a description of education reform in Texas.

- Conclusions and recommendations for improving the service that LEP students receive.

Duran, R.P. (1989). Assessment and instruction of at-risk Hispanic students. *Exceptional Children, 56(2),* 154-158.

This article presents an argument against current ways of using standardized tests with Hispanic bilingual students. The author describes two main limitations of these testing practices:

- Validity and reliability may be reduced for Hispanic students because of their lack of familiarity with test taking vocabulary, test taking strategies, and content.

- Test performance is not a good indicator of what needs to be learned next; it can only indicate what content students do not know.

The article discusses two recent developments in testing that would benefit Hispanic students. First, Tharp and Gallimore's work (1988) suggests that teaching only happens when the teacher helps a student accomplish a new task. This support from the teacher comes in the form of "useful hints and cues" while the student is doing the task. However, this method can not be used to teach students how to improve on standardized tests. Second, the procedure called "dynamic assessment" evaluates students' readiness to learn and teaches them cognitive skills. It includes a "test-train-test" cycle in which student performance is analyzed and students are given immediate feedback on how to improve their score. The author suggests that testing personnel diagnose students' learning potential and help teachers create more effective remediation for students based on the methods described in this paper.

Figueroa, R.A. (1990). Best practices in the assessment of bilingual children. In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology – II* (pp. 93-106). Washington, DC: National Association of School Psychologists.

This study, written for school psychologists, reviews the issues surrounding the testing of bilingual students. First, as a basis for discussing the best assessment practices for bilingual children, the author addresses four related topics:

- Legislation and standards for testing bilinguals.
- The nature of bilingualism.
- Education of bilingual students.
- Test bias and methods of non-biased assessments.

Second, the author makes recommendations in three major areas:

- Assessment of instructional programs that involve LEP students.
- Examination of the student's culture.
- Measurement of language proficiency.

Finally, the author rates current practices in the measurement of bilingual students' intelligence. The ratings given are, "not valid," "problematic," and "promising." In this last section, issues such as test translation, giving tests only in English, and the use of interpreters are discussed.

Fradd, S.H., McGee, P.L., & Wilen, D.K. (1994). *Instructional assessment: An integrative approach to evaluating student performance.* Reading, MA: Addison-Wesley.

This book provides information about "legally and educationally defensible assessment procedures for use with non-English language background (NELB) students." It explains how to choose appropriate assessments and how to use the information obtained from them to serve the needs of elementary and junior high LEP students. It gives guidelines for distinguishing students with second language acquisition problems from students with learning disabilities. Case studies describe different types of LEP students and show how school personnel used assessment to identify their needs. The book also discusses ways that teachers can become advocates for LEP students and how to develop a broad approach to assessment for use in planning, teaching, and monitoring students' progress.

Gandara, P., & Merino, B. (1993). Measuring the outcome of LEP programs: Test scores, exit rates, and other mythological data. *Educational Evaluation and Policy Analysis, 15(3),* 320-338.

This study reports on test data, reclassification of LEP students, and exit rates from LEP programs. It was part of a larger study by the California Legislature to compare the effectiveness of language learning programs for LEP students. The research questions that this part of the study set out to answer were: What is the relative rate of academic progress and second language acquisition across program models, and which model is the most successful at exiting students into mainstream classes? (p. 320) The results of the study indicated that there was a conflict between theory and actual school practice. As a result, the research question could not be answered. Inherent in the study were theoretical problems that schools faced in the areas of service (who should be served and for what period of time, and how do you define "proficient"?) and testing (what do you test and what standard do you use?). In addition, there were also several problems as a result of the national education context in which there is a lack of a standard definition for LEP and limits on accurately measuring the skills of LEP students. A lack of government funding for LEP student programs makes it difficult for schools with large LEP populations to do regular testing. For these reasons, schools did not often follow testing policies even though they indicated that they did. Sometimes LEP students are kept in ESL programs until teachers feel they are ready and are never tested when they exit. Sometimes students are permanently kept in ESL programs because the teachers feel that is the best place for them, and sometimes they are exited when they are still LEP. In addition, data were often collected inaccurately and were not comparable across programs. The researchers stress that some of these decisions were made for practical reasons. Based on the results of their study, several recommendations are made:

- Create a national definition of LEP with specific criteria so that the same student can be assessed in different places with similar results.
- Collect student data from performance-based tests so that LEP students do not have to be tested separately.
- Have a national effort to develop an accurate language proficiency test and an academic achievement test so that test data can be compared across the country.
- Do not tie funding for LEP students to exit rates.

Garcia, G.E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly 26(4)*, 371-392.

This paper reports on the results of a study comparing the reading achievement of bilingual Hispanic students and monolingual English speaking students on standardized reading tests. One hundred and four students in 5th and 6th grade were given a prior knowledge test to see whether they were familiar with the topics being tested, a vocabulary test of both general vocabulary and specific test vocabulary, and a reading comprehension test. A subset of these students was interviewed after the reading comprehension test to obtain data on how they responded to the test. The results of the study indicate that Hispanic students had lower reading achievement scores and that the following factors influenced their scores:

- *Vocabulary familiarity*. Hispanic students knew much less of the specific vocabulary in the reading passages than the scores on their vocabulary tests indicated. There were three major problems with vocabulary: (1) the students had trouble applying a word that they knew in one context to a different context. This was especially true when the word in question appeared as a different part of speech than the usage they were familiar with; (2) when the students read a word that looked like a Spanish word, they tended to apply the Spanish meaning of the word even if it did not fit the context; and (3) the students could not recognize material in the question and the correct answer choice that were paraphrases of each other.

- *Prior knowledge*. The students did not always activate the proper schemata because the passage did not define the key terms that were required to understand it. They also did not always know enough about a concept to rule out answer choices with incorrect information.

- *Test-wiseness*. There were five specific problems in this area: (1) students did not understand the textual relationships in the passage; (2) they could not locate answers that were specifically stated in the passage; (3) they did not recognize paraphrased material in the question and the answer choices; (4) they had difficulty answering questions that required them to draw conclusions; and (5) if they did not know the correct answer, they chose any answer that included words mentioned in the question even if that answer was incorrect.

- *Time*. Hispanic students generally needed more time than the Anglo students. The amount of time used was directly related to the test-taking strategies students utilized.

The researchers believed strongly that the interview process gave them more insight into the Hispanic students' reading abilities than the test did because students could say why they chose incorrect answers. The authors state that the unknown vocabulary in the questions and answer choices was the major linguistic factor that created difficulty for the students. When the test was translated into Spanish, more of the students had higher scores. However, it was difficult to account for the lack of prior knowledge about topics and the difficulty Hispanic students had making inferences since they were fluent English speakers not receiving any second language learning services. More research is called for to determine whether Hispanic students are receiving different instruction from Anglo students and whether this instruction causes the lack of skills in the two areas.

Gerken, K.C. (1978). Performance of Mexican American children on intelligence tests. *Exceptional Children, 44,* 438-443.

This article reports on a study of the relationships of the type of intelligence test (verbal or non-verbal), examiner ethnic group membership, and the language dominance of the student being tested to the performance of Mexican-American children on these tests. The article starts by discussing the shortcomings of previous studies, and then describes the research design. The study found that Mexican children performed better on non-verbal tests and that students who were dominant in Spanish had lower scores than English-dominant children. However, the impact of the examiner's ethnic group appeared to be much less significant than researchers anticipated. The author states that the study results indicate that Mexican American children are not "intellectually deficient" in comparison to other children and so their low rate of school success must be related to other factors.

Hamayan, E.V. & Damico, J.S. (Eds.). (1991). *Limiting bias in the assessment of bilingual students.* Austin, TX: Pro-Ed.

This book covers the field of bilingual special education assessment, but it has many chapters that are relevant to bilingual assessment in general. A few of the most relevant chapters have been annotated here. See separate entries for Oller & Damico (1991), Chamberlain & Medinos-Landurand (1991), and Cloud (1991). A chapter by Hamayan and Damico (1991), relating to language proficiency, will be included in a future bibliography.

Ingels, S.J. (1993). *Strategies for including all students in national and state assessments: Lessons from a national longitudinal study.* Paper presented at the National Conference on Large-Scale Assessment of the Council of Chief State School Officers, Albuquerque, NM.

This paper reports on the participation of students with disabilities and students with limited-English proficiency in the National Education Longitudinal Study of 1988 (NELS:88). A major finding was that 72% of the LEP students and 52% of special education students who were excluded could have participated. This lack of inclusion can have serious implications for making policy decisions for the general student population and for LEP students, and students with disabilities in particular. It can also affect the way that the data are used and the conclusions that can be drawn from them. The author makes recommendations on ways to include more LEP and special education students in longitudinal assessments and explains the importance of doing so. He then applies the analysis of the NELS:88 study and the recommendations made to the upcoming NCES Early Childhood Longitudinal Study (ECLS-K) and discusses the implications for this important longitudinal assessment.

Lacelle-Peterson, M. & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review, 64(1),* 55-75.

This article argues that educators must seriously consider the effects of educational reform, in particular testing reform, on LEP students. In the past, educators and policymakers have assumed that whatever testing policies work for monolingual, native-English-speaking students will also work for LEP students once they have learned enough English. If LEP students are truly expected to achieve high academic standards, an assessment system must be designed that measures the linguistic and academic abilities of these students. The authors believe that the current practice of waiting until an LEP student has reached a specific English proficiency level before teaching them academic content is harming LEP students. Frequently, by the time the student has the required level of English proficiency, there is too little time left to learn the academic content needed to meet high academic standards and go on to further education. Bilingual programs for teaching academic content play an important role in helping LEP students learn the content skills. For these reasons, the authors recommend an assessment system that gives information about the development of the student's proficiency in both academic and social English, the student's first language if it is appropriate (i.e., if instruction has been given in the first language), and the knowledge of academic content. Since second language learners differ in the speed at which they acquire spoken and written English, LEP students should be allowed to complete assessments in their strongest mode of expression and in their strongest language. Such a test would give more accurate results than a standardized

test which may only measure a student's English ability instead of giving a picture of their cognitive functioning and content knowledge. This test would give a complete picture of student progress over time instead of performance at an isolated point in time. Currently, such a system does not exist and many schools exempt LEP students from standardized assessments of educational progress as a way of recognizing that the standardized tests are inappropriate for them. The authors commend the effort to deal with inappropriate assessment systems but state that exempting students allows school districts to ignore the educational progress of LEP students. Instead, the authors call for an improved assessment system that has been designed to meet the needs of both native English speakers and non-native English speakers. The article concludes by giving five guidelines for establishing an equitable assessment system.

Lam, T.C.M. (1993). Testability: A critical issue in testing language minority students with standardized achievement tests. *Measurement and Evaluation in Counseling and Development, 26*, 179-191.

This article addresses several key issues concerning the misapplication of standardized achievement testing with LEP students. Five topics are discussed:

- Testability as redefined with regard to LEP students.
- Testability of LEP students as described in reference to five assumptions about the test taker's characteristics.
- Problems with exempting LEP students on the basis of language proficiency.
- Three alternative methods for establishing testability criteria for LEP students.
- Strategies for enhancing the applicability of standardized achievement testing of LEP students.

Recommendations are given for establishing appropriate guidelines for exempting LEP students from standardized testing and for developing special testing for these students. Further research in the area of testability of LEP students is recommended.

Lam, T.C.M. & Gordon, W.I. (1992, Winter). *State policies for standardized achievement testing of limited English proficient students, 11 (4),* 18-20.

This article reports the results of a questionnaire sent to the Department of Education in all 50 states and Washington, D.C. The questionnaire asked about testing policies for LEP students. The results of the survey indicate that there is "a general lack of state policies and guidelines regarding the standardized achievement testing of LEP students" (p. 20). Based on these results the authors suggest that current testing practices vary widely across states and policies are made somewhat arbitrarily. Because of the inconsistency of policies and procedures, bias is "probably introduced" into standardized achievement testing. The authors call for more research to be done to evaluate testing practices which can reduce the effect of limited English proficiency on test scores. In general, the authors state that the more support a state gave to bilingual/ESL education, the more likely the state was to have well developed policies for the testing of LEP students. It is recommended that State Departments of Education increase their support for such programs in order to create a better testing climate for LEP students.

National Academy of Education. (1996). *Quality and utility: The 1994 Trial State Assessment in Reading: the fourth report of the National Academy of Education Panel on the Evaluation of the NAEP Trail State Assessment: 1994 Trail State Assessment in Reading* (pp. 53-73). Stanford, CA: National Academy of Education.

Chapter Four of this document specifically relates to the assessment of students with disabilities and LEP students in the 1994 NAEP Trial State Assessment. The researchers found that there was a general lack of agreement on the interpretation of the published NAEP guidelines. For this reason, in 1994, more than 50% of the LEP students in the districts examined were excluded from taking the NAEP test. Educators usually cited a lack of oral language ability and a low level of reading proficiency as the reasons for exclusion of these students. However, when the NAE conducted its own research, it was found that many of the excluded students had been in the district longer than the specified amount of time needed to be considered for exclusion. In addition, more than 78% of the excluded students were found to be capable of taking the NAEP assessment. The researchers emphasize that in order to obtain high quality data on the general educational progress of all students, LEP students must be included in the assessment to the greatest extent possible. The students may not do well on the test, but they need to be included so that the data are accurate and can be compared across districts and across states. Based on their research findings, the NAE panel makes the following recommendations for NAEP:

- Continue to encourage greater participation of LEP students.

- Continue research on adaptations and accommodations for LEP students (research has shown that when accommodations are allowed, LEP students are more likely to be included in the assessment).
- Report the results for LEP students along with those for other students (do not disaggregate the scores for LEP students).
- Develop a test that can give more accurate measurements over a broader range of student proficiency levels.

Based on the results of the research discussed in this report, changes were recommended for the 1996 NAEP test to ensure greater participation of LEP students.

National Center for Education Statistics. (1996). *Increasing the inclusion of students with disabilities and limited English proficient students in NAEP* (Pre-Publication Copy). Washington, D.C.: National Center for Education Statistics.

This document describes changes made in NAEP to allow for greater inclusion of students with disabilities and LEP students. NAEP policies and procedures prior to 1996 are described first with a chart of inclusion rates for 1992-1994. These items are contrasted with new policies and procedures in place for the 1996 NAEP. The only new accommodation made for the 1996 NAEP was the allowance of Spanish-English bilingual test booklets for mathematics, or Spanish-only test booklets. Finally, a description is given of current studies in NAEP that focus on the inclusion of students with disabilities or LEP students.

National Center for Education Statistics. (1996). *Proceedings of the conference on inclusion guidelines and accommodations for limited English proficient students in the National Assessment of Educational Progress: December 5-6, 1994* (NCES 96-86(1)). Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement.

This report summarizes the results of a meeting held to provide assistance to NCES staff on the inclusion guidelines and accommodations for LEP students in the National Assessment of Educational Progress (NAEP). Participants recommended:

- Creating a standard definition of LEP so that school districts across the country interpret the term in the same way.
- Basing participation decisions on students' literacy in English (only those with

"sufficient" literacy levels should be tested using the unmodified English version of the test).

- Translating the test for students who do not have enough literacy in English (start with a Spanish translation).
- Simplifying the English vocabulary and/or allowing modifications to the test (i.e., giving answers on tape instead of writing them, having a glossary at the end of the test, etc.).
- Creating LEP scoring rubrics that consider the students' linguistic and cultural background.
- Giving districts a simple decision-making tree to make consistent participation decisions.
- Reporting the test results in three ways: (1) all students who took the test, (2) LEP students, and (3) non LEP students (it may also be desirable to report on LEP students who took the test with modifications and detail what the modifications were).
- Reporting where a student might score if he or she had taken the test (background information about the student could be included such as length of time studying English) and a portfolio of work could be kept.
- Monitoring the participation of LEP students carefully so that maximum participation is assured.

Navarrete, C., & Gustkee, C. (1996). *A guide to performance assessment for linguistically diverse students*. Albuquerque, NM: Evaluation Assistance Center – Western Region.

This document discusses the ways in which performance assessments can be used to systematically document the educational progress of linguistically diverse students. First, it defines the term "linguistically diverse," and then it goes on to give: the definition of a performance assessment and issues that need to be considered in designing such an assessment for a linguistically diverse student; a framework for selecting and designing a performance assessment (six basic elements of a performance assessment are discussed and a review of current literature is provided for each element to show the implications for linguistically diverse

students); ways to meaningfully present the data obtained from a performance assessment; and an appendix with sample assessments and a scale for rating them.

Oller, J.W. Jr., & Damico, J.S. (1991). Theoretical considerations in the assessment of LEP students. In Hamayan, E.V., & Damico, J.S. (Eds.). *Limiting bias in the assessment of bilingual students* (pp. 77-110). Austin, TX: Pro-Ed.

This chapter from a book on assessment of bilingual students discusses language theories and the way that they relate to assessment. The authors state that it is important for educators to have a theory that they ascribe to because the theory will influence the choice of tests and how the test is administered. The chapter begins by reviewing Cronbach's (1970) idea of a theoretical "construct" and the way a construct both guides and is shaped by practice. The theory of "language proficiency" is a construct that underlies the assessment of bilingual students. As a part of a review of empirical research on language assessment procedures and tests (no "clearly defined" literature was available on language proficiency), the authors discuss three paradigms of testing that have been influential in the past 30 years:

- *Discrete point testing*, which led to isolating components of language proficiency and testing them independently of each other.
- *Integrative testing*, which led to proficiency being assessed in a general context of language ability.
- *Pragmatic testing*, which combined pieces from discrete point and integrative testing.

The authors introduce a new "hierarchical model of representational capacities," based on the work of C.S. Peirce, which they believe will address some of the problematic issues that arose in the three paradigms. This new model has three main implications for testing bilingual students:

- Testing should assess performance in a rich context.
- Testing should support students and identify their "optimal capabilities" instead of focusing on their disabilities.
- Assessment should involve more than just one test.

The model supports Krashen's (1985) input hypothesis (i.e., students need to receive input from many different sources and in different modes), Cummins & Mulcahey's (1978)

threshold hypothesis (i.e., a certain level of proficiency is needed in both the first and second languages in order to benefit from instruction in either language), and a modified version of Cummins' BICS/CALP theory (i.e., the acquisition of interpersonal-communicative language takes place earlier than the acquisition of cognitive-academic language).

Olmedo, E.L. (1981). Testing linguistic minorities. *American Psychologist, 36(10),* 1078-1085.

This article addresses the topic of testing linguistic minority students. People from linguistic minority groups believe that standardized testing limits their access to full participation in the "social, political, and economic benefits" of society. First, the author describes the linguistic minority population in the U.S. Next, the "social, educational, and occupational status" of this group is discussed. If testing is to be meaningful, it must recognize the relationship of educational opportunity, language, and culture to status issues. Language ability affects standardized test scores, which in turn affects access to opportunities. Third, the literature relating to testing of linguistic minorities is reviewed. Since the 1960s, the idea of cultural pluralism has had a significant impact on testing. Finally, the author discusses some important conceptual and operational issues that relate to testing linguistic minorities:

- Effect of bilingualism on test performance and how test scores reflect language development.
- Effect on test scores of acculturation to majority norms.
- Cross-cultural generalization of cognitive constructs.
- Appropriate determination of language dominance for test administration.
- Test translation and the problems associated with it.
- Test administrators and the effect they have on the outcome of assessments.

Olson, J., & Goldstein, A. (1996). *Increasing the inclusion of students with disabilities and limited English proficient students in NAEP* (Focus on NAEP Report – Pre-Publication Copy). Washington, D.C.: National Center for Education Statistics.

This report summarizes information about the continuing development and implementation of the National Assessment of Educational Progress (NAEP). First, it discusses the history of

inclusion policies for students with disabilities and students who are limited English proficient. Then it explains the 1996 policy that emphasizes the inclusion of as many students with disabilities and LEP students as possible. For LEP students, NAEP math test booklets were available in Spanish and Spanish-English versions because the majority of LEP students came from Spanish-speaking backgrounds. NAEP gathered data on the impact of the new inclusion criteria and will publish the results in early 1997. A number of other studies relating to inclusion of students with disabilities or students with limited English proficiency are in progress. These studies address the areas of:

- Scaling issues.
- Reporting issues.
- Appropriateness of inclusion criteria for students with disabilities and LEP students.
- Construct validity of the assessment for students with disabilities and LEP students.
- Language complexity issues.
- Inclusion procedures.

O'Malley, J.M., & Pierce, L.V. (1994). State assessment policies, practices, and language minority students. *Educational Assessment, 2(3),* 213-255.

This study looked at two questions:

- What assessment policies are used by states for identification, placement, and reclassification of English language learners (ELLs) in special language programs designed to meet their education needs?
- What assessment policies are used by states in statewide testing programs, and what are the adaptations, if any, for ELL students?

The sample consisted of state education agencies (SEAs) in the eastern half of the U.S. (n=34 states and territories). A survey questionnaire was the data collection method (mailed in June, 1991) with a response rate of 100% (with follow-up call and personal contacts). States were divided into high-impact and low-impact categories depending on the number of ELL students. Both categories varied greatly in recommended or required policies for the identification and

placement of ELL students. The use of Home Language Survey (HLS) and an English language proficiency test for screening and identification purposes was the most consistent variable in both categories. Results indicated that the impact of the assessment programs depended on how the states used the results of the assessment. Six recommendations for SEAs are discussed.

Rivera, C., & Vincent, C. (1996). *High school graduation testing: Policies and practices in the assessment of limited English proficient students.* Paper presented at the Council of Chief State School Officers, Phoenix, AZ. Arlington, VA: Center for Equity and Excellence in Education.

This paper reports the results of a 1994 survey of state assessment directors who were asked about their policies for LEP students taking state assessments. Results indicate that in 1993-94, seventeen states required students to pass content area tests to get a high school diploma. These states dealt with LEP students in the following ways:

- First administration of the test was delayed until students had more English proficiency.
- Modifications were used.
- The test was available in the student's first language.
- An alternative assessment was given.

The authors found that the four methods listed above made the test accessible to only a limited number of LEP students, and this lack of access may have future legal implications. The authors make the following recommendations to states who have or are developing graduation tests:

- Collect data on: (1) the impact of the test on certain populations (i.e., do they all have equal access to the content that is required in order to pass the test?); (2) the test performance patterns by race, ethnicity, socio-economic status, gender, and LEP classification; (3) the performance of test repeaters and the drop out rate for students who do not pass; and (4) the effectiveness of remediation/interventions for LEP students who do not pass.
- Use temporary exemptions from testing sparingly.
- Use modifications cautiously and have clear policies for when they can be used.

- Develop tests in native languages for students who have received enough content instruction in their first language. If the state has large numbers of students who share a first language, the development of a psychometrically equivalent test in that language is recommended.

- Use alternative assessments for LEP students who can not take other forms of the test. The alternative assessments should be aligned with instruction in order to be effective.

Saville-Troike, M. (91, Spring). Teaching and testing for academic achievement: The role of language development [On-line]. *NCBE Focus: Occasional Papers in Bilingual Education. No. (4).* Available: http://www. ncbe.gwu.edu/ncbepubs/focus/focus4.html

This article discusses some new theories in second language acquisition that may have an impact on assessment of LEP students. The author believes that the traditional focus on language proficiency and its relationship to academic achievement has been too narrow. First, the idea of positive transfer of both scripts for school and background knowledge is discussed. The author summarizes her research which shows that students who have completed two years of education in their native country generally succeed in U.S. schools despite their level of English language proficiency. Second, the "watering down" of curriculum for LEP students is described. Schools have a tendency to believe LEP students incapable of learning academic content if they do not have much proficiency in English. Sheltered English classes may not be helping LEP students because they teach down to the students' English level. Third, the author lists some other factors besides language proficiency that may combine to affect academic success. One must not separate language proficiency from these other factors and treat proficiency as the sole cause of failure to succeed. Finally, the issue of assessment is discussed. The author believes that "radical changes are needed in testing procedures and interpretation" and makes recommendations for testing based upon her research. One important testing issue is that knowledge of vocabulary seems to be more highly related to academic success than language proficiency.

Spencer, B. (1991). *Eligibility/exclusion issues in the 1990 Trial State Assessment.* Chicago, IL: Northwestern University, Methodology Research Center, National Opinion Research Center.

This report discusses issues related to groups of students who were exempted from taking the 1990 Trial State Assessment (TSA), which was given in 37 states, Washington, D.C. and two

U.S. territories. The design of the TSA allowed three groups of students to be exempt:

- Students in private schools.
- LEP students who had been in an English-speaking school less than two years and were judged incapable of taking the assessment.
- Special education students with IEPs who were judged incapable of taking the assessment.

The author conducted "illustrative sensitivity analyses" to give an estimate of how states' scores and rankings might have changed had these three categories of students been involved in the testing. The results of these analyses are discussed and recommendations are given for changing exemption policies on NAEP and reporting data more accurately in the future.

Wilen, D.K., & van Maanen Sweeting, C. (1986). Assessment of limited English proficient Hispanic students. *School Psychology Review, 15(1),* 59-75.

This article explores the assessment of LEP Hispanic children in American schools and provides information about "potential pitfalls." It also gives suggestions for the assessment of LEP students in general. The article begins with cautions for people who assess LEP Hispanic children because the results of assessment are affected by socio-cultural problems faced by many Hispanic immigrants to the U.S., the influence of the native culture, and the process of second language development. The article then gives recommendations for assessing Hispanic students. These recommendations include: obtaining background information about the student, giving language proficiency tests in both Spanish and English, giving a non-verbal intelligence test, using both formal and informal methods to gather information on academic ability, and, in cases of possible mental retardation, giving an adaptive behavior assessment. A bilingual team of school personnel is recommended when conducting assessments. By following the authors' recommendations, the assessment process should be fairer and more meaningful for Hispanic bilingual students.

Zehler, A.M., Hopstock, P.J., Fleischman, H.L., & Greniuk, C. (1994, March 28). An examination of assessment of limited English proficient students [On-line]. *Task order D070 Report*. Arlington, VA: Special Issues Analysis Center. Available: http://www.ncbe.gwu.edu/miscpubs/siac/lepasses.html

This document reviews issues related to the assessment of LEP students. First, three types of assessment are discussed:

- Assessment to promote high levels of achievement (part of education reform).
- Assessment to determine placement in a language program or eligibility for services.
- Assessment for school/program/district accountability.

Any discussion about assessment of LEP students needs to identify the type of assessment. Second, it is important to consider the purposes of assessment. An individual student may be assessed for placement in a program or to exit from that program, for language ability, for academic ability, and for some type of disability. A group of students may be assessed for progress as a group, for program effectiveness, or for accountability. The authors give a list of characteristics/components of each of the nine individual and group purposes listed, and they rate each characteristic in terms of its importance for each of the assessment purposes. For example, testing English language proficiency is rated as "very important" for identification, placement, language assessment, and review of placement. It is rated as "moderately important" for academic assessment. The authors include recommendations for educators based upon the characteristics that were discussed. Third, the results of a literature review on assessment practices of successful programs are described. The authors found articles giving recommendations for assessment practices with LEP students, but no empirical research studies. Because of the lack of available research, the authors reviewed "the applications from nine first year Title VII Academic Excellence Projects funded in 1993." Results indicated that few of the projects included assessment as a distinct component. Since LEP student assessment does take place, it is likely that project writers viewed assessment as separate from the effective instruction programs outlined in the applications. Fourth, because language proficiency assessments may be used for many different purposes, the following six language proficiency tests are reviewed and compared: (1) Idea Proficiency Test 1, (2) Language Assessment Scales 1C (Oral), (3) Language Assessment Battery 1A, (4) Bilingual Syntax Measure II, (5) Maculaitis Assessment Program, Level 2-3, and (6) Peabody Picture Vocabulary Test – Revised. The content and nature of test items are compared, and the administration procedures for each are discussed along with the theoretical bases of the tests. Issues of reliability and validity are also discussed. Fifth, criticisms and concerns in using standardized tests are listed.

In the sixth section, criticisms and concerns in using alternative assessments are listed. The authors appear to favor alternative assessments, but state that this type of test can be more demanding because they require LEP students to use more language. Finally, the article ends with conclusions and recommendations for the future of assessing LEP students.

General Testing Issues

Airasian, P.W. (1987). State mandated testing and educational reform: Context and consequences. *American Journal of Education, 95*, 393-413.

This article details recent social and educational changes that have influenced the educational testing process. In the past 20 years, four trends have affected the educational system:

- A greater tendency to diagnose societal problems and to look for ways of solving them.
- The belief that environment affects human behavior, and that the time to try to change behavior is when humans are young (policy makers try to change organizations in order to produce the desired behavior in people).
- The push for more equity and equal opportunity for underrepresented social groups (schools are now required to serve the needs of many groups of people who were not served before; this desire for equality has focused more attention on the outcomes of education).
- The changes in the American family (resulting in schools now having a greater role in socializing children without the supportive relationship with homes that once existed).

When these four trends are combined with the huge "expansion, centralization, and politicization" of the educational system, tests take on new roles. They are now used to monitor the entire educational system and to attest to individual performance within that system. The author discusses the characteristics of these new testing programs and the consequences of them. A set of proposals is given to provide a framework for understanding the new context of testing.

Bond, L.A., & King, D. (1995, November). *State high school graduation testing: Status and recommendations*. Oak Brook, IL: North Central Regional Educational Laboratory.

This paper focuses on the obstacles that states face in developing a graduation testing program. The obstacles relate to:

- Determining what knowledge should be tested and at what level.
- Determining a passing score.
- Guaranteeing that students have had a chance to learn the material before being tested on it.
- Assuring that a minimal competency test does not determine the cut off for learning and figuring out what to do about the students who do not pass the test.

The paper first looks at the history of state graduation testing programs from the late 1970s to the present. After the 1991 document *What Work Requires of Schools: Secretary's Commission on Achieving Necessary Skills* (SCANS), there has been a push to assess higher level skills and this push often conflicts with minimum competency testing. The authors state that the tension between assessing basic skills and assessing higher standards is a fundamental issue that states must deal with when constructing their tests. The paper then examines graduation tests in the 18 states that currently have them. No single testing model is used by every state. Every state uses norm-referenced, multiple-choice tests; writing assessments are common as well. However, testing occurs anywhere from 6th grade to 11th grade. Some states use minimum competency tests and some test higher standards. Because the states do not all follow the same model, and because the tension between minimum competency and higher standards testing is unresolved, it is difficult to determine whether the graduation testing program as a whole is successful. Based on the inconclusive results of reviewing state tests, the authors conclude by making several recommendations for improving the assessment design and implementation process.

Cuevas, J. (1996, January). *Educating limited-English proficient students* (PD-96-01). San Francisco, CA: WestEd.

This monograph describes program options and appropriate teaching approaches for ESL programs in elementary and secondary schools. The authors give a description of the LEP population in the U.S. and the educational rights of LEP students. Second, a discussion of

language competency, including Cummins' (1981) theory of cognitive/academic language proficiency, takes place. Third, ways of assessing LEP students' language proficiency and academic ability are listed. The author discusses both formal and informal assessments. Next, theories of second language acquisition are related to classroom teaching and testing. Attention is paid to Krashen's hypotheses on language learning. Effective strategies for teaching academic content to LEP students are presented. These strategies include: sheltered English, cooperative learning, thematic instruction, multicultural education and pluralistic education (teaching that includes multiple perspectives), and the language experience approach. Appropriate teaching methods are then described. The author reviews research that supports a move toward communicative, natural, and total physical response approaches instead of the traditional audio-lingual approach to language teaching. The history of English language learning programs in the U.S. is examined and a range of program options such as submersion, pull out, and structured immersion are discussed. ESL pull out classes are popular because they are less costly and easier to implement. A discussion of the plusses and minuses involved in each type of language program follows. The author supports bilingual education and believes that academic content classes should be taught in the student's first language when possible. Finally, the author describes schoolwide improvements and changes that can be made in order to support LEP students.

Gerken, K. (1990). Best practices in the academic assessment of secondary-age students. In A. Thomas & J. Grimes (Eds.). *Best practices in school psychology – II* (pp. 13-28). Washington, D.C.: National Association of School Psychologists.

This chapter discusses problems with both the current system of assessing secondary students and the remediation process when a student does not pass an assessment. A hypothetical case study of a student at risk is given to illustrate the shortcomings of traditional assessment practices and the ways in which a student can slip through the cracks because they do not fit into an easily identifiable category such as "learning disabled." These problems are partly related to a lack of adequate training for school psychologists in academic assessments for secondary students. They are also partly due to a need for improved use of the chosen assessment instead of continually searching for a newer and better one which does not exist. The chapter discusses ways to improve the screening and diagnosis processes, and then goes on to look at improving assessment in the content areas of reading, math, writing, and science. Examples of formal and informal measures are given for each content area.

Kean, T.H. (1991, April 24). Do we need a national achievement exam? Yes: To measure progress toward national goals. *Education Week, 28*, 36.

This commentary piece, written by the president of Drew University in Madison, New Jersey and the chairman of Educate America, supports a national achievement exam for high school students. The author believes that a test is needed to measure progress toward the national goals that have already been set, and lists five reasons why the test is needed:

- Students need to study harder and be held accountable for what they learn.
- Schools could be held accountable because all schools would report test data to the public.
- Students would be assessed on difficult content that they have studied instead of on minimum skills.
- Results would be clear indicators of what students have learned, and the public could understand the results.
- The test would be a high quality assessment used to measure critical thinking skills and it could, in turn, force school curricula and teaching practices to improve.

Linn, R.L., Baker, E.L., & Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*, 15-21.

In recent years there has been a growing interest in alternative forms of assessment. However, educators and researchers who have argued for alternative assessments have not given data showing that these assessments are more valid than norm-referenced, standardized tests. This paper lays out eight criteria for evaluating alternative assessments:

- What are the consequences of assessment? (e.g., Will teachers spend too much class time preparing for the test?).
- Is the test fair to students from different gender and racial/ethnic backgrounds?
- Is the test a reliable indicator of student performance on other tasks?
- Does the test measure higher order thinking skills? What types of responses are required?

- Is the test content up-to-date and high quality?
- Is there a broad representation of content coverage that fits the wide range of understandings in each field? (not all experts agree on what content should be assessed, so these different views should be accounted for on the test).
- Is the content meaningful and worth assessing?
- Is the assessment cost effective?

By using these criteria to evaluate a test, standardized tests and alternative tests can be compared to determine which type is more valid in a given situation.

Mehrens, W. (1993). *Issues and recommendations regarding implementation of high school graduation tests.* Regional Policy Information Center. Oak Brook, IL: North Central Regional Educational Laboratory.

This is the first in a series of policy papers related to high stakes student assessment programs. The paper describes trade-offs in policy making so that educators can make informed decisions. Two short articles introduce the paper:

- "High School Graduation Requirements in the North Central Region," edited by Roger Trent of the Ohio Department of Education, which lists what each state in the North Central Region required for graduation in 1993.
- An executive summary by Linda Ann Bond, Director of Assessment for NCREL Regional Policy Information Center.

The paper itself was originally a report written for the Michigan Department of Education by an Expert Panel of eight members who were asked to help Michigan implement its new high school graduation test. The report was broadened to be applicable to any state working on a graduation test. It describes issues that need to be resolved, gives recommendations, and presents a list of tasks to be performed along with a time line for completing them. The issues discussed relate to core curriculum/test specification, psychometrics, education, legal aspects, policy/administration, and human resources/financial resources.

Neill, M. (1991, April 24). Do we need a national achievement exam? No: It would damage, not improve, education. *Education Week, 36*, 28.

This commentary piece, written by the associate director of FairTest, the National Center for Fair & Open Testing, argues against a national achievement test for high school seniors. The author takes issue with the popular belief that assessment improves the quality of education and specifically argues against Thomas Kean's proposal for a testing program. Neill feels that a national, standardized test fails to address some of the critical issues such as fairness, rigid school leadership, and poor textbooks. Such a test could be used as a "gatekeeper" to separate out students by race and social status. Currently, the discussion of the exam is taking place before a set of content standards has been put in place and this is problematic. Instead of a standardized test, Kean proposes a performance-based program that would allow for individual differences in the students' background and would measure higher order thinking skills in a way that a standardized test could not. The country could mandate the use of this type of testing system without requiring the same test for all students. In this way, the test could be sensitive to needs of students in a given area, and content area reform could happen from the bottom up, instead of from the top down.

Nutall, E.V. (1987). Survey of current practices in the psychological assessment of limited-English-proficiency handicapped children. *Journal of School Psychology, 25*, 53-61.

The purpose of this study was to determine the assessment practices currently used for LEP children with disabilities mainstreamed into bilingual classrooms. The study looks at the types of tests used, the personnel involved in administering the assessments, and the problems encountered in assessing this group of children. At the time of the study, government officials were interested in this information for the purpose of determining funding priorities for research and training programs in bilingual education. Twenty-one local education agencies (LEAs) were included in the study sample. The major data collection methods were telephone and personal interviews (data were collected in 198[3]). Results indicated that the most frequently used testing approaches included the adaptation or translation of existing tests and the common cultural approach. The majority of LEAs reported shortages in bilingual assessment personnel. This shortage caused the LEAs to implement priorities for testing and to create lists of students waiting to be tested. Recommendations include the need to encourage bilingual students to enter into the field of school psychology, as well as the need for new approaches and testing procedures to improve the effectiveness of school psychologists who work with LEP students.

Phillips, S.E. (1994). *Legal implications of high-stakes assessment: What states should know*. Regional Policy Information Center Report. Oak Brook, IL: North Central Regional Educational Laboratory.

This report was written to "help state and national education policy makers avoid legal challenges to their student assessment programs." The author recommends that states consider the possible legal challenges to their assessment as they are creating the program, and she gives guidelines for developing defensible programs. The guidelines are based on past legal decisions and cover four main areas:

- Testing to award diplomas.
- Potential bias against historically disadvantaged groups.
- Testing accommodations for disabled persons with disabilities.
- Performance assessment issues.

"Each chapter describes relevant legal measurement, and policy issues; analyzes applicable federal statutes and case law; and presents recommendations for legal defensibility" (p. 2). LEP students are not specifically discussed, but much of the information is applicable to situations that involve these students.

Shinn, M.R., & Tindal, G.A. (1988). Using student performance data in academics: A pragmatic and defensible approach to non-discriminatory assessment. In R.L. Jones (Ed.). *Psychoeducational assessment of minority group children: A casebook* (pp. 383-407).

This paper begins by outlining the need for non-discriminatory assessments for minority and LEP students. Traditional norm-referenced tests are not often normed on a population similar to the ones from which the minority or LEP student comes. Therefore, the[illegible]e biased because they require certain types of cultural knowledge that the minority or LEP student may not have. The authors lists three current assessment practices that attempt to limit bias in norm-referenced testing used to identify special education students:

- Using non-verbal tests.
- Translating or interpreting a test into the student's first language.

- Using more than one model of a "non-discriminatory assessment" (e.g., culture free tests, culture fair tests, culture specific tests, and pluralistic assessment tests).

Each one of these practices is problematic, and the drawbacks to each are discussed. Then the authors describe an alternative assessment practice that is informal, less culturally biased, and obtains data on student performance which can be used to more easily improve instruction. This type of assessment involves the continuous collection of data using performance-based measures. Performance assessments for reading, spelling, and writing are described. Over a ten-week period, student performance on these measures was comparable to performance on a norm-referenced test and was less biased. In addition, the performance-based assessment demonstrated student improvement over time and was a more accurate indicator of whether the student needed special education services. Finally, three case studies of LEP/minority students are given to show how the continuous performance assessment helped to identify learning difficulties and plan remediation.

Walstad, W.B. (1984). Analyzing minimal competency test performance. *Journal of Educational Research, 77(5),* 261-266.

This article reports on a study of the impact of school district resources and policy changes on basic skills test performance. The authors believe that school officials may be able to improve education for basic skills if they know how school policies, student characteristics, and other community factors affect test performance. In this study, two years of data from the Missouri Basic Essential Skills Test (BEST) were used and the model chosen to represent the educational process was: "Basic Achievement (B) in a school district is a function of: the family background of students (F); the resource and organizational inputs provided by the school system (S); and, the student labor time (L) applied to education." (p. 262). After a lengthy discussion of the equation used in the study, the rationale for using this particular model of the educational process and the particular characteristics of the BEST test, the authors describe how they chose the values of each variable in the equation based on characteristics of the BEST test. The results of the study indicated that school policies (e.g., changing curriculum, training teachers, etc.) did not affect test performance except for the practice of pre-testing students the year before they took the BEST test to determine whether the students needed additional educational support. Pre-testing was found to significantly increase test scores and it had the added benefit of being relatively inexpensive and easy to implement. The results also indicated increased funds the district spent on each pupil did not produce consistently better test results. Instead, the size of the school district seemed to produce more significant results; bigger school districts were shown to have a higher percentage of students passing the BEST test. The author

recommends that as school districts look for ways to improve their test scores on minimal competency tests, they consider whether the strategies they choose to implement are effective.

Wilde, J., & Sockey, S. (1995). *Evaluation handbook.* Albuquerque, NM: Evaluation Assistance Center – Western Region.

This handbook offers suggestions on how to conduct good evaluations and gives information and guidelines aimed at program directors, staff, and evaluators of IASA's Title VII bilingual programs. It has six sections:

- An overview with some definitions of evaluation design terms and standards for evaluation that were developed by the Joint Committee on Standards for Educational Evaluation in 1981.
- Background information about and definitions of evaluation, assessment, and analytic techniques. Descriptions of types of evaluations and guidelines for managing evaluations.
- Information about planning evaluations: developing timetables, writing measurable objectives, selecting assessments, selecting scoring methods, etc.
- Implementation (specific needs of Title I and Title VII students are addressed here).
- Guidelines for interpreting data, presenting results, making recommendations, and writing reports.
- Appendices with more detailed information (e.g., frameworks for program evaluation, methods of determining goals and prioritizing them, Title VII Reporting Procedures, etc.).

Psychometric Issues in Assessment

Bracken, B.A., & Barona, A. (1991). State of the art procedures for translating, validating and using psychoeducational tests in cross-cultural assessment. *School Psychology International, 12*, 119-132.

This paper explores the best procedures for translating, validating, and using tests across

languages and cultures. In the past, test translation was not generally accepted because there was no well established procedure for doing it. Factors such as difficult test instructions, and the differing importance of concepts in the various languages stood in the way of providing the translations. The authors review the work of other researchers, such as Werner and Campbell (1970), and Bracken (1984, 1987, 1990), who have begun to establish guidelines for the "appropriate use and interpretation of translated tests in cross-cultural assessment." After a brief discussion of basic considerations in test translation, the authors outline an eight-step procedure for translating and validating tests. The steps are:

- Make an identical translation from the source language to the target language.
- Have a new translator, who is unfamiliar with the original test, translate the test from the target language back to the source language.
- Repeat this process until the discrepancies between the two versions have been reduced as much as possible.
- Have a bilingual review committee examine the translated version to determine whether parenthetical words, vocabulary changes, etc. are needed and whether items on the test are appropriate for students from minority cultures.
- Do a pilot test with a bilingual psychological examiner and examinees from different social, economic, and geographical backgrounds.
- Do a more extensive field testing.
- Develop local, regional, or national norms of the scale depending on how the test will be used.
- Validate the translated test.

The authors emphasize that in addition to examining the test itself for bias, the context of the test should be examined as well. Issues such as the ability of the examiner to speak both languages, the student's fluency and educational history in both languages, and the level of difficulty of vocabulary in the translated test (the process of translation might cause the vocabulary to be harder or easier on the translated version than on the original) should be considered. Also important are issues such as the immigration history of the student, and the beliefs, customs, values, and degree of acculturation of that student. The examiner's beliefs about immigrants and minorities play a role as well.

Cline, T. (1993). Educational assessment of bilingual pupils: Getting the context right. *Educational and Child Psychology, 10(4),* 59-68.

This article looks at the importance of evaluating the context of assessments given to bilingual students before the process of assessment is examined. In order to decide whether the assessment is reliable and valid, three types of context must be examined: social, intellectual (the working assumptions of the people involved in the testing), and educational.

Examining the social context involves looking at a student's reasons for being bilingual and the relative status of each of the student's languages. It also includes consideration of social pressures that the child's ethnic/linguistic community is facing. Examining the intellectual context involves looking at the assumptions that educators hold about four main issues: bilingual language proficiency, second language learning, assessment, and learning difficulties. The educator's personal biases can influence the testing process and the way that results are used and interpreted. Examining the educational context involves looking at the current education reforms, the way that test data are collected and used, and the environment of the school (e.g., Are there other speakers of the student's language in the school? Are the materials available in the student's first language? Are there specific school policies for bilingual students?). After all of these context-related factors have been examined, then educators can look at the test itself. Two main questions should be asked about the test: (1) Who is involved in the assessment – parents, teachers, students, support staff, et al.?; and (2) What is being assessed? The author recommends that data on LEP student achievement be gathered from more than one source and that data are kept on individual student progress.

Dorans, N.J., Schmitt, A.P., & Bleistein, C.A. (1992). The standardization approach to assessing comprehensive differential item functioning. *Journal of Educational Measurement, 29(24),* 309-319.

This article describes the standardization approach to assessing differential item functioning (DIF) on the Scholastic Aptitude Test (SAT). A test item is said to show DIF when one group of test takers scores lower or higher on that item than other groups do. The authors use the term "comprehensive differential item functioning (Cdif)" to refer to the standardized approach to assessing DIF. The authors discuss a way to calculate DIF, a way to calculate Cdif, and then compare these to other approaches such as the log-linear approach for assessing "differential distractor functioning" (Green, Crone and Folk, 1989) and the IRT approach to differential alternative functioning (Thisen, et al., 1992). Data from an SAT test are given to show how the

standardization approach to Cdif can be used to show differential speededness (different response rates between a reference group and another group). In the data given, it can be seen that Asian American test takers do not show evidence of differential speededness, but both black and Hispanic test takers do. The researchers found that when black and Hispanic students took a version of the SAT in which the order of the test sections was changed (a common occurrence in administering the SAT to maintain test security) the amount of differential speededness also changed. When a section with more items and more difficult items in the beginning came first, the black and Hispanic students showed more differential speededness than they did when the section in question was second. This phenomenon is important because differential speededness may result in high DIF for items which come at the end of a test section and which are not completed by all students. The implication for test administrators and graders is that DIF may be artificially induced by the format of the test and this factor needs to be taken into consideration when deciding whether to use certain test items.

Fuchs, D. & Fuchs, L. (1989). Effects of examiner familiarity on Black, Caucasian, and Hispanic children: A meta-analysis. *Exceptional Children, 55(4),* 303-308.

This paper presents an argument for examining the context of an assessment to determine whether that assessment is biased. Charges of bias are often subjective and it is important to look at factors related to the administration of the test when considering issues of fairness. Previous research on one of these factors, examiner familiarity, has shown that children with disabilities perform significantly higher when they are tested by someone they know. The research reported in this paper was done to determine whether the same finding would hold true for minority students. The study involved a review of available literature on minority students and examiner familiarity. Twenty two studies were found on the topic, and of these, thirteen studies were evaluated for quality. The findings from these studies indicate that black and Hispanic children did score significantly higher when they had a familiar examiner. However, the research does not conclusively indicate whether race was the important factor in the improved test scores with a familiar examiner. Socio-economic status may have been the most important factor, and this was not controlled in the studies. In addition, it is not possible to show that the performance of the minority students was representative of the performance of all minority students, so the data are not generalizable. The authors call for more research on the effects of race and examiner familiarity to validate their findings.

Hills, J.R., Subhiya, R.G., & Hirsch, T.M. (1988). Equating minimum-competency tests: Comparisons of methods. *Journal of Educational Measurement, 25(3),* 221-231.

Typically, inconsistent results have been found among the common IRT equating methods used to develop equivalent tests. This study examines the effects of five different equating procedures on cut-points and the number of anchor items needed to develop equivalent tests. The test selected was Florida's Statewide Student Assessment Test, Part II (SSAT-II). The 1984 and 1986 test administrations served as the metric of comparison. The five equating methods used were: (1) The Linear Method (LINEAR), (2) The Rasch Model (RASCH), (3) The Three-Parameter IRT: Concurrent Method (IRTCON), (4) The Three-Parameter IRT: Fixed Method (IRTFIX), and (5) The Three-Parameter IRT: Formula Method (IRTFOR).

Results indicate differential effects among the equating procedures across content areas. Among the five procedures, results indicate anomalous behavior for the IRTFIX method in low score regions on the Communication section. For example, using the IRTFIX method would have resulted in 2,700 fewer students passing the mathematics portion of the test because the cut-point for passing would have been higher. In the mid-score distribution for Mathematics, IRTFOR procedures produce anomalous results, while the other IRT methods above the 16th percentile produce scores within one raw score point. All anchor tests (25- to 10-item) were found to be within a quarter of a raw score point from the standard anchor test. Only the 5-item anchor test was found not to be an adequate sample for equating assessment measures. Overall, the IRTCON procedure appeared to allow for the fewest anchor items. Fewer anchor items are often a benefit to test developers and reduces the number of items needing to be repeated across test years, thus enhancing test security. This superior and significant finding for IRTCON applies only to the number of anchor items needed and does not apply to other methods of equating.

Huynh, H. (1990). Error rates in competency testing when test retaking is permitted. *Journal of Educational Statistics, 15(1),* 39-52.

This document is a heady formula-laden paper discussing two types of statistical estimates of error rates, Type I, declaring a child competent who truly is not; Type II, failing to declare a child competent who is truly deserving which can occur when a student is allowed multiple trials on a mastery test. The paper examines two procedures (beta-binomial and Rasch Model) to estimate Type I and Type II errors for students who are permitted to retake a test that he or she initially failed. Theoretically, in competency testing where a student is permitted to repeatedly retake a test, the probability of a false negative is zero. Establishing the probability of false positive error rates is, however, much more challenging. Traditionally, when choosing between beta-binomial and Rasch models it is assumed that the Rasch model estimates are more accurate due to the fact that they take into account more data. To examine the reality of this basic assumption this study analyzed six data sets and compared two statistical procedures. The

findings suggest that beta-binomial estimates tended to be slightly larger than Rasch estimates for false negative errors (Type II). This should not be a large issue given that repeated retakes of a mastery test reduce this error rate to zero. On the other hand, beta-binomial error rates were much larger than Rasch model error rates for false positives (i.e., declaring a child as competent who is truly not). While psychometrically the Rasch model appears to be the statistic of choice, it is a more difficult statistic to obtain and requires more comprehensive data. When only means and standard deviations are available, beta-binomial estimates appear to be sufficient for most practical considerations

Pearson, B.Z. (1993). Predictive validity of the Scholastic Aptitude Test (SAT) for Hispanic bilingual students. *Hispanic Journal of Behavioral Sciences, 15(3),* 342-356.

This study reevaluates the predictive validity of SAT scores by looking at the college success of Hispanic bilingual students who had low scores on the test. The participants were those entering the University of Miami in the fall of 1988. The students were divided into two groups: Hispanics and non-Hispanic whites. The University of Miami GPA (after 4 semesters), SAT-Verbal, and SAT-Math scores were compared for the Hispanic and non-Hispanic students. Results indicated that Hispanic students had a significantly lower mean SAT score than non-Hispanic white students. However, the two groups had equivalent college grades.

Schmitt, A.P., & Dorans, N.J. (1990). Differential item functioning for minority examinees on the SAT. *Journal of Educational Measurement, 27(1),* 67-81.

This paper was written by two researchers from ETS and it summarizes recent findings about differential item functioning (DIF) for minority students who took the SAT exam. An item is said to show DIF if "the probability of correctly answering the item is lower for examinees from one group than for examinees of equal ability from another group or groups" (p. 68). After discussing the definition and explaining how it is calculated, the authors discuss DIF studies for Asian Americans, Hispanics and blacks who took the SAT exam. It was found that Asian Americans who did not speak English as their best language had more items with negative DIF than did any of the minority students for whom English was their best language. This finding held true even on the math section where many problems had "a high 'verbal load.'" For this reason, Asian Americans who did not speak English as their best language were dropped from the study and only students who reported English as their best language were included. It was also found that on sentence completion and reading tasks, items with

content that was familiar to black and Hispanic students had positive DIF and seemed to be easier for them (e.g., in an analogy question, a reference was made to a "dashiki," which is an African word for a piece of clothing; African American students had positive DIF on this type of item). All the minority students had negative DIF for items that had homographs in them. All the minority students in the study completed test items at a slower rate and were said to show "differential speededness." This different rate of test taking tended to show negative DIF for test items at the end of a section because minority students often did not finish a test section. The researchers recommend further research in the areas of the relationship between "vertical associations" (words in the stem and the distractor are related even though the distractor is not the correct answer to the stem) and negative DIF.

Miscellaneous

Anstrom, K. (1996, Summer). Defining the limited-English proficient student population [On-line]. *Directions in Language and Education.* Available National Clearinghouse of Bilingual Education: http://www.ncbe. gwu.edu/ncbepubs/directions/dir19.html

This paper, written by the NCBE information analyst, discusses the variety of approaches used to measure the LEP student population in the United States. Before measurement of the LEP population can take place, there must be a working definition of "LEP." The definitions that are used vary a great deal, and this has a significant impact on the estimate of the LEP population obtained by different measurements. To illustrate the problem, Anstrom gives the federal definition of LEP and contrasts it with state definitions from New York, California, and Texas and Cheung's (1994) definition from "The Feasibility of Collecting Comparable National Statistics about Students with Limited English Proficiency: A Final Report of the LEP Student Counts Study." The advantages and disadvantages of having a standardized definition of LEP are discussed. While having a national definition would help to give a more accurate count of the LEP population, a way to more accurately identify students in need of service and some common terminology for discussing the needs of this population, such a definition could also be problematic. A standardized definition might be biased in favor of some groups and would increase the responsibilities of and the costs for the school districts that need to do additional testing. The last part of the article is an annotated bibliography of estimates of the LEP student population.

Council of the Great City Schools. (1996). *Becoming the best: Standards and assessment development in the great city schools* [On-line]. Washington, D.C.: Council of the Great City Schools. Available: http://www.cgcs.org/reports/1996/standard

This research report summarizes the results of a survey done by the Council of the Great City Schools (CGCS) in the spring of 1995. The forty-seven districts that belong to the CGCS were surveyed and asked to describe the process of developing and implementing higher standards and assessments. Directors of research and evaluation were asked to complete the surveys, but sometimes other school staff filled them out. Several themes arose from the survey results:

- Schools are attempting to raise the standards and are moving in step with state and national reform efforts.
- School districts are using state and national resources such as standards work that has already been done (they believe that the information they receive from their state has been very useful, and they request more help in the area of technical assistance relating to assessment).
- Schools are writing standards for core content area subjects and are using these standards to develop assessments, change curriculum and instruction, and improve programs.
- Districts are concerned about funding for the implementation of the new standards (because of limited funding, they are waiting to see how states respond to Goals 2000 before they do too much themselves).

Districts are also concerned about assessment and are, again, waiting to see what states require before they develop their own requirements.

Del Vecchio, A., & Guerrero, M. (1995, December). *Handbook of English language proficiency tests*. Albuquerque, NM: Evaluation Assistance Center – Western Region, New Mexico Highlands University.

This handbook provides educators with a resource for evaluating and choosing a commercial, standardized English language proficiency test. The handbook does not recommend a test; it lists characteristics of each so that educators can make informed choices. Before reviewing the tests, the authors give background on the legal mandate to test students' oral language proficiency before placing them in an LEP class. Theories of language proficiency and

definitions of LEP are reviewed because there is controversy in both of these areas. Because proficiency can vary in different contexts, the authors emphasize the importance of choosing a test that asks students to do tasks similar to what they need to do in class. A discussion of types of language proficiency tests illustrates the importance of aligning one's theory of second language proficiency with the test that is chosen. Possible types of tests are discrete point (e.g., a phoneme discrimination test or multiple choice vocabulary test), integrative or holistic (using more than one skill at once such as listening to a story and retelling it), and pragmatic (linking the test with the student's experience; "real life"). A caution is included about the limitations of language proficiency tests and the different types of classifications they can produce. The five tests reviewed are: (1) Basic Inventory of Natural Language (BINL), (2) Bilingual Syntax Measure I and II (BSM I & II), (3) Idea Proficiency Tests (IPT), (4) Language Assessment Scales (LAS), and (5) Woodcock Munoz Language Survey.

Issues such as cost, reliability, validity, administration time, scoring, etc. are discussed. At the end of each critique is a list of resources for more information on the specific test that was reviewed.

Devine, J. (1988). The relationship between general language competence and second language reading proficiency: Implications for teaching. In R.C. Anderson & P.D. Pearson (Eds.). *Interactive approaches to second language reading* (pp. 260-277). New York: Cambridge University Press.

This chapter is a literature review of research on the relationship of second language competence and second language reading proficiency. According to the author, most researchers and educators believe a certain level of second language competence is necessary before a student can be an effective reader in that language. The article starts with the history of research in the field and leads into three research questions for this study:

- Does limited proficiency in a language restrict general reading ability in that language?
- Does limited proficiency in a foreign language restrict readers from using very specific types of textual information, such as discourse constraints, when reading in that language?
- Is there a threshold of linguistic competence that readers must reach before they can read successfully in a second language? How can that threshold, if it exists, be defined? (p. 263)

The results of the literature review indicate that there is a significant relationship between low reading ability and low language proficiency. In addition, poor second language readers are not able to use context clues and cohesive devices to determine relationships and to fully understand the reading. There is less evidence for a "linguistic threshold" and the authors call for more research to be done in this area. The article finishes with suggestions for teaching reading based upon the findings of the literature review.

Guerrero, M. & Del Vecchio, A. (1996, March). *Handbook of Spanish language proficiency tests*. Albuquerque, NM: Evaluation Assistance Center – Western Region, New Mexico Highlands University.

This handbook is similar in structure to the *Handbook of English Language Proficiency Tests* by Del Vecchio and Guerrero. The introduction discusses the Spanish-speaking population of the United States and the numbers of Spanish-speaking students. A widespread need to assess the Spanish language proficiency of school children is indicated. Information about legal mandates for proficiency testing and definitions of language proficiency is also included. The handbook does not recommend a test; it lists characteristics of each so that educators can make informed choices. Before reviewing the tests, the authors give background on the legal mandate to test students' oral language proficiency before placing them in an LEP class. Theories of language proficiency and definitions of LEP are reviewed because there is controversy in both of these areas. Because proficiency can vary in different contexts, the authors emphasize the importance of choosing a test that asks students to do tasks similar to what they need to do in class. A discussion of types of language proficiency tests illustrates the importance of aligning one's theory of second language proficiency with the test that is chosen. Possible types of tests are discrete point (e.g., a phoneme discrimination test or multiple choice vocabulary test), integrative or holistic (using more than one skill at once such as listening to a story and retelling it), and pragmatic (linking the test with the student's experience; "real life"). A caution is included about the limitations of Spanish language proficiency tests and the different populations of Spanish speakers. Finally, five tests reviewed are: (1) Basic Inventory of Natural Language (BINL), (2) Bilingual Syntax Measure I and II (BSM I & II), (3) Spanish Idea Proficiency Tests (IPT), (4) Language Assessment Scales (LAS), and (5) Woodcock Munoz Language Survey.

Issues such as cost, reliability, validity, administration time, scoring, etc. are discussed. At the end of each critique is a list of resources for more information on the specific test that was reviewed.

McDill, E., Natriello, G., & Pallas, A. (1985). Raising standards and retaining students: The impact of the reform recommendations on potential dropouts. *Review of Educational Research, 55(4),* 415-433.

This article discusses the possibility of an increased high school drop out rate due to the new school reform policies. First, the article reviews research on the drop out rate and why students drop out of school. The research shows that most students drop out for family and economic reasons. Holding a regular job while in school is a strong predictor of dropping out. Second, the article summarizes reports that recommend increasing the educational standards in secondary schools to improve student achievement. Third, the article weighs potential positive and negative results of higher standards. Research shows that some students will work harder and achieve more with higher standards, especially if they receive support from the school when they have learning difficulties. However, the authors believe that higher standards will mean that schools require a strictly academic course of study that does not suit all students. Options like vocational education will not be offered. This academic course will require more time from the students, and potential drop outs may not have this extra time because of their responsibilities outside of school. The end result may be a higher level of student frustration and failure without many opportunities for remediation. The authors emphasize that there is no research to support these conclusions; they are based on speculation. Finally, the authors suggest seven areas in which additional research is needed:

- Systematic evaluation of educational policies relating to higher standards.
- Precise data on school attendance and drop outs.
- Use of the "full enrollment model."
- Specific school conditions that result in higher achievement under the new standards.
- Additional resources needed by students in order to meet the standards.
- Ways in which responsibilities outside of school are passed on to students.
- Consequences of raising standards.

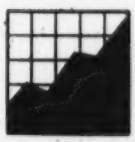
United States General Accounting Office. (1994, January). *Limited English proficiency: A growing and costly educational challenge facing many school districts.* Report to the Chairman, Committee on Labor and Human Resources U.S. Senate. Washington, D.C.: United States General Accounting Office. GAO/HEHS-94-38.

This report was written at the request of the Chair of the Committee on Labor and Human Resources in preparation for the reauthorization of federal elementary and secondary education acts. The report addresses five questions:

- What are the characteristics of LEP students nationally and in selected districts?
- What are the challenges districts face in educating these students?
- How do selected districts with LEP students from linguistically diverse backgrounds educate these students, including the extent to which academic subjects are taught in the students' native languages?
- What approaches have been identified as promising when diversity of language spoken by students makes native language instruction difficult?
- Do key federal programs targeted to LEP students provide the types of support districts need to implement programs to serve these students?

The report focuses on the availability of bilingual education. The data for the report were obtained by (1) analyzing 1980 and 1990 census data, (2) visiting districts in California, Massachusetts, New York, and Texas to find out how they were educating LEP students, (3) reviewing the literature, interviewing experts, and visiting districts to determine successful teaching practices for LEP students, and (4) interviewing Department of Education officials and other experts.

The report finds that there is limited support for LEP students to study content area subjects such as math and social studies. Bilingual education is a promising way to educate these students, but it is only available to some students. It can not be offered to all because of the number of different language groups represented in one school district. There is a shortage of bilingual teachers and of bilingual curriculum materials. In the absence of these teachers and materials, there are promising teaching approaches that monolingual classroom teachers can use. However, these approaches take time and training on the part of the teacher and are costly to implement. Federal funding for LEP students has dropped off in recent years and districts do not have enough money to meet the demand for services.



NATIONAL CENTER ON EDUCATIONAL OUTCOMES

University of Minnesota • 350 Elliott Hall
75 East River Road • Minneapolis, MN 55455
612/626-1530 • Fax 612/624-0879
http://www.coled.umn.edu/NCEO

The College of Education & Human Development

UNIVERSITY OF MINNESOTA

BEST COPY AVAILABLE